DNA Conformation Dynamics and Human Diseases

Boian S. Alexandrov, T-1; Alan R. Bishop, PADSTE; Luba Kalaydjieva, Western Australian Institute for Medical Research; I. S. Kohane, Childrens Hospital, Harvard Medical School; Kim Ø. Rasmussen, T-5; Anny Usheva, Beth Israel Deaconess Medical Center; V. I. Valtchinov, Harvard Medical School

The inherent thermal motion that a DNA molecule is subjected to in a living cell makes it possible for the two strands of the molecule to locally and spontaneously open and close. This dynamic effect is referred to as “DNA breathing” because it induces local transient separation of complementary base pairs. Numerous prior studies that combine modeling and experiments have firmly established a connection between this localized breathing dynamic of DNA and functional characteristics of mammalian transcriptional promoters. Here we highlight two recent studies that utilized the relationship between the breathing dynamics of DNA and biological function to elucidate some of the mechanisms of pathogenesis and progression in cognitive deficit schizophrenia, Friedreich’s ataxia, Huntington’s disease, and Fragile X-syndrome.

In living cells DNA molecules are subjected to inherent thermal motion, which makes it possible for each DNA base pair, or for a number of consecutive base pairs, to spontaneously open and re-close. This effect is referred to as “DNA breathing” because it induces local transient openings of the double helix. Numerous studies have shown that this local conformational dynamic of DNA is important for genomic functions such as transcription and replication. Previous research studies that combine modeling and experiments [1–4] have firmly established a correlation between the local breathing dynamics of DNA (also known as bubbles) and functional characteristics of mammalian transcriptional promoters. The results suggest that the breathing dynamics of DNA can more generally be correlated with regulatory genomic regions, and as such can be important for biological function. It is, however, uncertain whether breathing dynamics might be a determining factor in pathogenesis or in a disease state. Therefore, it is important to examine whether genomic mutations that compromise local bubble patterns may enhance the risk of a particular disease. Such genomic mutations are not necessarily located in regulatory genomic regions. Their disruption of biological function (e.g., suppressed transcription) may also be caused by long-range changes of the physicochemical or structural properties of the double helix. Importantly, various changes of the nucleotide sequence can lead to a similar impairment in conformational breathing dynamics. Thus, seemingly inconspicuous variations in DNA sequence may result in disease onset or progression. Here, we highlight two recent studies that utilized the relationship between the breathing dynamics of DNA and biological function to elucidate some of the mechanisms of pathogenesis and progression in cognitive deficit schizophrenia [5], Friedreich’s ataxia, Huntington’s disease, and Fragile X-syndrome [6].

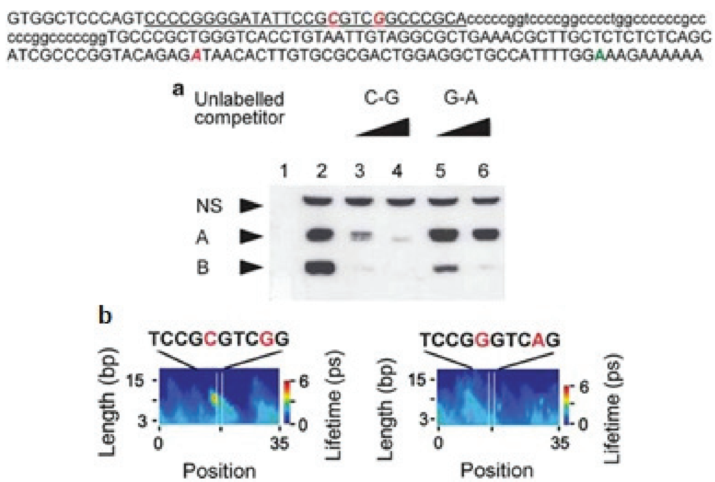

Fig. 1. DNA breathing dynamics and gel-shift experiments. The examined DNA sequence is displayed at the top with the SNPs in red and the transcription start site in green. (a) Gel-shift assay with nuclear extracts. Unlabeled oligonucleotides used in the self- and cross-competition assays are shown on top, with triangles indicating increasing concentrations (20–100-fold molar excess) of the competitor. Lanes: 1, no nuclear extract; 2, no competitor; 3 and 4, self-competitor; 5 and 6, cross-competitor. Bands: NS, nonspecific; A and B, specific, with decreasing intensity in the self-competition experiment. Preferential binding of the protein complex to the C-G sequence is suggested by the inability of the G-A oligonucleotide to out-compete the labeled C-G probe. (b) Langevin dynamics simulations of DNA breathing. Strong breathing activity can be seen for a sequence containing the C-G sequence (left), in contrast to the poor breathing activity of the G-A sequence (right) [5].

Cognitive deficit schizophrenia. In a recent study published in *Molecular Psychiatry* [5], we studied the effect of single-base-pair variations (single nucleotide polymorphisms [SNP]), which have been implicated in cognitive deficit in schizophrenia, on DNA dynamics and gene expression. Specifically, we report experimental evidence of gene expression combined with modeling of DNA breathing dynamics that points to two adjacent promoter SNPs as the functional variants associated with schizophrenia. Our modeling revealed significant

For more information contact Boian S. Alexandrov at boian@lanl.gov

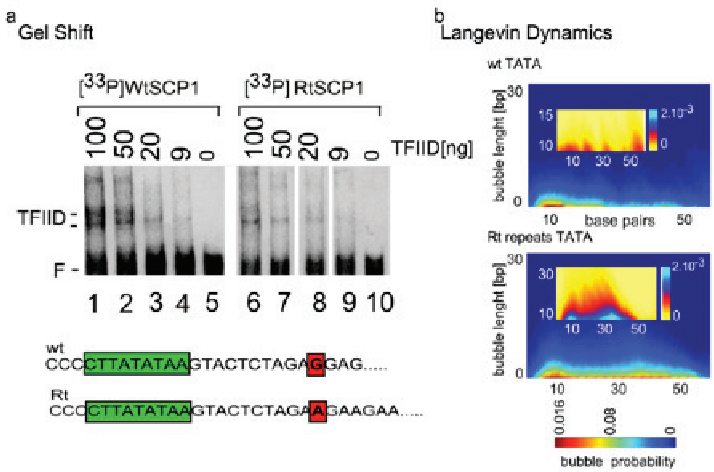

Fig. 2. Effect of TRSs on protein DNA binding and local DNA dynamics. DNA wild type (wt) and mutated (Rt) sequences displayed at the bottom. Protein binds to the segment in green. The Rt has been modified to contain 15 (GAA.TTC) repeats downstream of the TATA box as indicated. a) Band shift titration reactions received double-stranded oligonucleotides containing the shown sequences. The reactions in lanes 1–4 and 6–9 received different amounts of purified protein as indicated at the top of the lanes. The reactions in lanes 5 and 10 did not receive protein. The positions of the free DNA (F), and transcription factor (TFIID) are indicated on the left. b) Langevin dynamics simulations of the DNA breathing dynamics in Wt and Rt sequences [6].

sequence effects of the SNPs on the local breathing dynamics of DNA (Fig. 1). In particular, strong breathing activity was predicted for a sequence containing the C–G sequence (Fig. 1, left) in contrast to the poor breathing potential of the G–A sequence (Fig. 1, right), which causes a preferential binding of the protein complex to the C–G sequence and changes the gene expression. Gel shift and transcription experiments confirmed the conclusions from our simulations (Fig. 1) [5].

Trinucleotide repeat disorders. Friedreich's ataxia, Huntington's disease, and Fragile X-syndrome are examples of genetic diseases associated with the expansion of trinucleotide repeat sequences (TRS). We have found that in TRSs, the local breathing spectrum is altered by the increased symmetry that results from the repeats. It is commonly accepted that the TRS expansion is caused by the formation of non-B-DNA structures that disrupt normal cellular processes. In addition, it is believed that the initiation of non-B-DNA structures can be caused or enhanced by transient DNA openings. Hence, we simulated opening profiles of TRSs with various lengths and GC-contents. Our simulations suggested that large transient DNA bubbles can be formed not only in A/T-rich sequences (e.g., (GAA.TTC)n, Friedreich's ataxia), but also in sequences with a relatively high G/C-content, such as (CTG. CAG)n in Huntington's and (CGG.CCG)n in Fragile X-syndrome. Our simulation data implicate the transient DNA openings as a common factor in the trinucleotide repeat disorders. Further, we found that TRS sequences tend to open more cooperatively as the length of the repeats increase. This collective behavior leads to abrupt bubble creation that significantly enhances the presence of intermediate DNA bubble states, as compared with sequences with a random nucleotide composition or with much shorter repeat tracts. In summary, we have demonstrated that TRS expansion in the disease-related (longer) sequences leads to enhanced coherent DNA openings – that is, enhanced local strand separations – when compared with the "healthy" sequences with a low number of repeats. Finally, we provided an experimental example of how such enhanced TRS bubble spectra can interfere with protein-DNA interactions, and thus with biological function. We demonstrated that the flankings composed of repeats have a profound effect on the spectrum of the TATA-box DNA-breathing dynamic, which could explain the lost TFIID-TATA binding (Fig. 2). This novel correlation between the transient bubble spectrum and repeats expansions in the individual genomes suggests local DNA dynamics as an "epigenetic" determinant that plays an important role in the TRS genetic diseases.

[1] Bishop, A.R. et al., *Springer Series in Materials Science* **148,** Kakeshita, T., et al. Eds., Springer, New York (2012).
[2] Alexandrov, B.S. et al., *PLoS Comput Biol* **5,** e1000313; PMID: 19282962 (2009).
[3] Choi, C. et al., *Biophys J* **95,** 597; PMID: 18390611 (2008).
[4] Alexandrov, B.S. et al., *Nucleic Acids Res* **37,** 2405; PMID: 19264801 (2009).
[5] Jablensky, A. et al., *Mol Psychiatr,* doi: 10.1038/mp.2011.129; PMID: 21968932 (2011).
[6] Alexandrov, B.S. et al., *PLoS One* **6,** e19800; PMID: 21625483 (2011).

Funding Acknowledgments
American Recovery and Reinvestment Act-NIH; National Library of Medicine; National Health and Medical Research Council of Australia; New Atlantic Technology Group, LLC; LANL Directed Research and Development Program